

Erin R. Lemelin. Digital Archival Collections Hosted on Flickr: Institutional Experience Outside of the Commons. A Master's Paper for the M.S. in L.S degree. November, 2010. 41 pages. Advisor: Jeffrey Pomerantz

This study describes an online survey of institutions which host digital archival collections on Flickr outside of the Flickr Commons. The survey was conducted to determine the experience of these institutions with Flickr in terms of usage statistics, user interaction, and assessment. The results from the survey were also used to determine factors which contribute to the success of a Flickr collection.

The majority of the eighty-nine institutions which responded to the survey reported satisfaction with Flickr and stated that they had achieved their goals. However, many respondents reported low usage or did not know the usage, which indicates weak evaluation. Factor analysis revealed that age, metadata creation, enabled social media tools, and number of staff members with time dedicated to the collection all related to the success of these collections. Recommendations for increasing usage are included.

Headings:

Flickr (Web site)

Virtual Library

Special Collections

Internet Surveys

DIGITAL ARCHIVAL COLLECTIONS HOSTED ON FLICKR:
INSTITUTIONAL EXPERIENCE OUTSIDE OF THE COMMONS

by
Erin R. Lemelin

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Library Science.

Chapel Hill, North Carolina

November 2010

Approved by

Jeffrey Pomerantz

Introduction

As the Web becomes ubiquitous and Web browsers become the major interface for using applications and data, users expect more and more content to be available online remotely. Librarians and archivists responded to this expectation by offering more resources which could be accessed through the Web. One of the widespread efforts which began because of this shift is the digitization of archival materials. Archival materials in particular are candidates for digitization because they are often fragile and valuable, which can make them difficult for users to access. These materials are also usually rare, one-of-a-kind items which, through digitization, can be viewed remotely and concurrently by multiple users rather than by a single user with physical access. As the number of digitization projects has increased, platforms have been created specifically for the display of these materials. CONTENTdm and Omeka are two examples of such platforms. They both allow institutions to create digital collections with Dublin Core metadata standards, customizable presentation, and self-containment. One of the alternatives which has a very low cost (unlike CONTENTdm) and minimal setup and maintenance (unlike Omeka), is Flickr. Flickr was not created to serve libraries and archives, so it lacks many of the functions which librarians and archivists would prefer to have, especially concerning metadata. Information professionals still saw potential use for the tool as a host for their digitized content and as a way to bring their archival materials to the user community.

One of the most successful digital archival collections on Flickr is the Flickr Commons. The Library of Congress established the Flickr Commons as a Web space for digitized visual materials with no known copyright restrictions. Any institution may apply to become a member, and the Commons currently consists of forty-six members. Users of the Commons are then encouraged to engage with the materials and create metadata with the various tools available (comments, notes, and tags). However, the process that institutions must go through to become members is time consuming, and the backlog of requests is so large that the Library of Congress is not accepting new members for the rest of 2010 (Flickr from Yahoo, 2010a). This delay is one of the reasons why some institutions have not joined the Commons. Other reasons include concerns over the “no known copyright restriction” and the freedom given to users to create metadata. Many archival collections are still under copyright constraints and therefore cannot be added to the Commons. Also, librarians and archivists have traditionally had tight control over the descriptive information about materials and have used controlled vocabularies. If librarians or archivists are uncomfortable with the freedom users have in the Commons, they could choose to have an independent Flickr account.

One of the disadvantages for institutions which do not participate in the Commons is the loss of some online exposure and user awareness of their collections. Flickr promotes the Commons through a number of access points and advertisements on the Flickr Web site. The Commons is also attached to some very large, prominent institutions, such as the Library of Congress and the New York Public Library, which adds to its publicity. Digital archival collections which are hosted only through accounts independent of the Commons do not receive this extra promotion within Flickr.

That is not to say that a collection hosted outside of the Flickr Commons will not still receive high usage. Many factors come into the “success” of a digital archival collection which is hosted on the Web. “Success” in the context of a Flickr collection is defined here as a combination of factors such as a high amount of views, comments, tags, notes, favorites, and more. As the data reported in this paper shows, most institutions have created Flickr collections in order to increase user awareness and interest in the collections. The quantifiable evidence of that use is partly in these tools and counters. Other manifestations of increased user awareness may be an influx of reference questions related to the collection and traffic to the institution’s Web site from the Flickr page.

The point at which low user awareness becomes high user awareness is more difficult to determine. The audiences for archival collections vary, so the expected user interest correspondingly varies. Ultimately, each institution has different goals for a Flickr project, and each institution evaluates the project’s success according to those goals. This paper will examine both raw numbers and institutional assessment in determining the “success” of a project.

As of yet, the institutions with digital archival collections on Flickr outside of the Commons have not yet been studied. Therefore, the purpose of this research is to discover:

1a. What are the experiences of institutions which host digital archival collections on Flickr outside of the Commons in terms of usage statistics, goal achievement, and increased user awareness?

1b. How do the experiences of the institutions in 1a differ (or not) from the experiences of institutions with collections in the Flickr Commons?

2. What are the factors which contribute to the “success” of a digital archival collection hosted on Flickr outside of the Commons?

Literature Review

Since Flickr is a relatively new tool on the Web, only a few studies have been conducted about it. The tagging habits of users on Flickr have been a topic study for some researchers (Angus, Thelwall, & Stuart, 2008; Rafferty & Hilderley, 2007; Stvila & Jorgensen, 2009). A paper which is directly related to this research is Jason Vaughan's survey of members of the Flickr Commons. In his survey, Vaughan collected information such as usage statistics, institutional satisfaction, and specifics on social interactions. Vaughan's study necessarily left out the digital archival collections outside of the Flickr Commons. His survey also does not attempt to determine factors which contribute to the success of a collection's online presence.

One of the other relevant papers is a case study concerning the Alcuin Society and its process and experience with Flickr (Saunders, 2008). Their experience overall was very positive: "At least one of the visitors to our Web site decided to blog about us. Online photographs showing people at our events and showcasing amazing, rare, and unique books tend to promote themselves. Without question, it pays to know what your audience is interested in" (Saunders, 2008, p. 307). Since this paper was a case study, however, its description of the Society's experience cannot be powerfully generalized. The paper does end with some implementation steps for using Flickr as a host, but these steps are a "how-to" guide, not a list of recommendations (Saunders, 2008, p. 307-8).

Another related paper is Mary Samouelian's study of Web 2.0 implementation in digital archival collections. Her article does not concern Flickr specifically, but since Flickr utilizes Web 2.0 technology, her findings do apply. ("Web 2.0 technology" is here defined as tools which allow social interaction and user generated content on the Web, such as comments, tags, favorites, and notes.) Fifty-seven percent of archivists surveyed in Mary Samouelian's study noted that "promotion of collections" was one of the reasons why they added Web 2.0 tools (2009, p. 63). Twenty-nine percent of respondents indicated that "participation from patrons" and "sharing content with potential users" was a reason for Web 2.0 implementation (Samouelian, 2009, p. 63). Finally, fourteen percent said that "staying current with our users" was part of the impetus for using Web 2.0 technology (Samouelian, 2009, p. 63). Later in her article, Samouelian shares that "[f]ive of the respondents (71%) answered that feedback from their patrons has been positive" (2009, p. 66). Samouelian's study, though it was conducted quite recently in 2009, contains a very small sample of digital archives with Web 2.0 technology. Her findings, like the Alcuin Society case study, are more difficult to generalize to the entire population of digital archives.

In the previously mentioned survey of institutions in the Flickr Commons, Jason Vaughan found results that were very similar to Samouelian's in terms of the goals for Flickr projects: "All but one respondent (94 percent) ranked 'expose collections to a broader audience/facilitate discovery of our materials' as a very important reason in joining The Commons; with the final respondent ranking this as 'important'" (Vaughan,

2010, p. 189). Each of the institutions also felt that they had achieved these goals concerning impact through the Commons (Vaughan, 2010, p. 198).

The user community's involvement in the Commons has been highly praised (Kalfatovic, 2008; Vaughan, 2010). The institutions involved are also some of the most powerful, significant forces in library science, such as the Library of Congress, the New York Public Library, and the Smithsonian (Flickr from Yahoo, 2010b). However, a comprehensive survey of the various institutions which have digital archival collections in Flickr outside of the Commons has never been done.

Methodology

The methodology for this study consisted of an online survey. The data from the survey is then reported and compared to the data Vaughan collected. Finally, factor analysis is conducted on certain variables against indicators of success, which will be discussed below.

Since no comprehensive list exists which contains the names of all institutions with digital archival collections on Flickr, the population had to be imperfectly estimated. In order to be included in the population estimate the digital collection must:

1. Be hosted on Flickr
2. Not be included in the Flickr Commons
3. Be associated with an institution or organization (non-profit or for-profit), not an individual
4. Contain at least fifty archival images

In this paper “archival” collections do not include contemporary, born digital photo collections. If a collection online contained both archival photos and born digital photos, the collection size for that institution included only the archival photo collection(s), which were in separate Flickr sets or collections. As long as at least fifty archival photos were included, the institution was considered part of the population estimate.

The population estimate was determined by searching the Flickr Web site through its native search engine for usernames containing the word “archive,” “special collection,” “library,” or “museum.” The resulting collections were viewed, and the researcher determined whether the collection qualified for inclusion using the previously mentioned criteria. If the collection qualified, then an associated e-mail address was harvested, where possible. Not all Flickr profiles have associated e-mail addresses, so when necessary, the researcher collected contact information from the institution’s home Web site. Whenever possible, the researcher collected the e-mail address of the staff member most likely in charge of the Flickr collection. This determination was made by job title. At smaller institutions, there was often only one individual associated with archival materials, so that staff member was contacted. When the institution was larger, there was generally a Head of Digital Collections and that individual was contacted. Occasionally the researcher discovered that only a generic e-mail address was available or that contact was only possible through a Web form. In those cases, the researcher used that contact information and requested the staff member who received the e-mail to either forward the survey solicitation to the appropriate staff member or inform the researcher of the contact information for the appropriate staff member. 207 contacts were collected, 18 of which received the solicitation through a Web form.

Data Collection Instrument. The contacts received an e-mail with an invitation to anonymously take an online survey (Appendix). The survey is based heavily on Jason Vaughan’s survey of the members of the Flickr Commons in 2009. Many of the questions have been slightly rephrased in order to apply to institutions which are not members of

the Commons. Most of the questions in Vaughan's survey referred to the "Flickr Commons," and in this survey they refer to "Flickr" in general.

The e-mail solicitation for the survey was sent between October 4, 2010 and October 6, 2010. The recipients received two reminders. The survey closed October 23, 2010.

Factor Analysis. Once the data was collected, certain factors were analyzed for association with various indicators of success. The factors that were chosen are as follows:

- age of the collection online
- time spent creating metadata
- number of staff involved
- institution type
- amount of social media features enabled
- whether or not the materials are online elsewhere

Indicators of success are as follows:

- average number of views per month
- total number of views
- percent of items which have received at least one tag, or at least one note, or at least one tag
- increased traffic to the home Web site attributed to Flickr
- time spent answering reference questions about materials in the Flickr collection
- amount of social media features used

- whether or not the respondent feels the goals have been met

The chi-square, t-test, and logical regression were used to determine the associations, if any, between these factors and success indicators.

Survey Results

The survey received 89 responses. Since $n=207$, the survey received a 43% response rate. Respondents were not *required* to answer any of the questions, however, so the response rate for each question varies.

Background Information. This section of the survey determined the history of the institution's Flickr collection. The age of the collections varied from one or two months to more than five years old, with the average being 21.45 months. The most important reason reported for using Flickr was to "expose collections to a broader audience / facilitate discovery of our materials," which had a mean rating of 3.67 out of 4 and most often received the rating "Very important reason." The second most highly rated reason was to "help advertise / provide a link to our institution," which had a mean rating of 3.05 out of 4. The third most highly rated reason was to "utilize Web 2.0 features to engage user involvement / discussion," which had a mean rating of 2.77 out of 4. "Desire to join the company of other prestigious institutions that were already a part of Flickr" was the least important, with a mean rating of 1.85 and the fewest "Very important reason" ratings. Most respondents (65%) stated that the reason why they chose to make these particular digital objects available online was because "the photo is regionally important." The second most chosen reason (47%) was because "the photo is part of a popular collection." The least chosen reason (23%) was that the materials were already online (see Table 1 below).

Table 1.

Question	Not applicable / not a consideration	Slightly important reason	Important reason	Very important reason	Responses	Mean
Expose collections to a broader audience / Facilitate discovery of our materials	1	4	14	57	76	3.67
Utilize Web 2.0 features to engage user involvement / discussion	9	21	22	22	74	2.77
We didn't already have these photos online and Flickr was a good system to initially publish them	22	12	15	25	74	2.58
It would help advertise / provide a link to our institution	4	14	31	26	75	3.05
Desire to join the company of other prestigious institutions that were already a part of Flickr	35	23	8	8	74	1.85

Staff Involvement. This section established the staff time spent on managing the Flickr collection. The number of staff working on the Flickr collections varied from 1 to 7 people, with the average being 2.18 staff members. 25% stated that only one staff

member maintained the Flickr collection. Respondents reported that staff time was spent most on the “selection of photos / collections with the direct intent that they will be incorporated into Flickr” and “uploading photos into Flickr.” Both receive on average 1 hour or less per week of staff time. The main task that respondents wrote-in for “Other” was the process of scanning materials. This task receives on average between 1.1 and 5 hours per week of staff time.

Social Interactions. This section concerned the types of social tools enabled and used in the collection. “User generated comments,” “bookmarking photos as ‘favorites,’” and “the incorporation of photos into viewing / discussion / subject groups set up by other Flickr members” are the features which are the most often both enabled and used in the collection. “User generated tags,” “user generated notes,” and “‘Blog this’ functionality” are the features which most often have been enabled but not used. “‘Blog this’ functionality” and “user generated tags” are the feature most often not enabled at all (see Table 2 below).

Table 2.

Question	Yes, this feature has been used by some users	No, this feature is enabled but hasn't been used by any users at all	Not enabled	Responses
User generated tags	17	25	13	55
User generated comments	42	13	4	59
User generated notes	18	28	9	55
“Blog This” functionality	11	20	10	41
Bookmarking of photos as “favorites”	43	10	3	56
Incorporation of photos into viewing / discussion / subject groups set up by other Flickr members	35	8	6	49

15% of respondents had tags disabled, 4% disabled comments, 10% disabled notes, 11% disabled ‘Blog this’ functionality, 3% disabled bookmarking, and 7% disabled incorporation into sets by other Flickr members. Many respondents either did not respond, did not know the statistics, or did not have access to statistics which would tell them the percentages for user interaction with their items (46% of respondents for tags, 39% of respondents for comments, and 47% of respondents for notes). An average of approximately 5.6% of these institutions’ photos received at least one comment, 2.2% received at least one note, and 3.3% received at least one tag.

Statistics. This section of the survey determined some of the general Flickr collection characteristics. The respondents reported the approximate number of views since going live, and the results varied from under 100 to over 800,000. The average total number of views over all collections is around 70,000. The reported approximate average amount of views per month varied from under 10 to 60,000, with the average over all

collections being around 4,000. Respondents also reported approximately what percentage of the materials online on Flickr were available elsewhere online. 21% stated that the photos in the collection were not available elsewhere; 32.5% stated that between 0.01-50% were available elsewhere; 6% stated that between 51-99% were available online elsewhere; and 7% stated that all of the photos were available elsewhere. When asked if they noticed an increase in traffic to their home Web site which they attribute to the Flickr collection, 35% said they had and 65% said they had not.

Assessment. This section of the survey determined how the respondent felt about the success of the Flickr collection. 80% reported that they felt they had met their initial institutional goals for the collection, and 20% did not. In terms of impact and popularity, 31% said that they met the majority of their goals, 65% said they met some of their goals, and 5% said they had not met any of their goals. Two respondents said they did not plan on continuing to use Flickr, and two respondents stated that they were not sure. One respondent said they planned to migrate the collection to other more robust services, and it appears that he/she would phase out the Flickr collection. Most respondents recommended Flickr, but some added caveats about cost, copyright concerns, and the need to prepare and plan properly. Interestingly, the cost of Flickr was positive or negative for different respondents:

- “It's a cost effective way for a smaller library to make its photo collections available and is easy to use.”
- “Yes, it is a very good resource for a very reasonable price. Especially helpful for small libraries with limited resources and budgets.”

- “[F]or small and medium-sized institutions like us, yes, because it is more affordable and generally more user-friendly than commercial [digital asset management systems] like contentDM ...”
- “But Pro costs extra.”
- “There is an ongoing cost for Pro Accounts which might change or be difficult to uphold.”
- “There is a limit as to how many pictures can go up [without] being charged.”

Overall, though, the respondents felt that the \$24.95 annual fee for a Pro account was a very reasonable and manageable cost for the service (Flickr from Yahoo, 2010c). In contrast, CONTENTdm costs thousands of dollars annually (Bond, 2006, p. 64; ILLINET, 2010, p. 8); and Omeka, while it is free in terms of software, must be locally hosted and managed. If an institution chooses to use Omeka, it must have the staff and technical resources to setup and maintain the framework.

Another concern that the respondents often mentioned was the challenge of handling the images’ copyright status carefully:

- “The major challenge we have encountered is selecting items for which copyright is not an issue. I think that is something that institutions need to be prepared to address.”
- “They should [be] aware of rights [management].”
- “... consider copyright (I put ‘all rights reserved’ on ours)”
- “Image rights, of course, must be considered [carefully] before posting and institutional control of the image.”

- “[T]hey should be aware of copyright issues specifically”
- “... be aware of privacy settings and copyright issues. Some artists do not want their images on Flickr.”

Not all archival materials are free of copyright restrictions, so institutions need to be aware of the copyright status of their materials. It might not be possible to digitize and allow online access to some materials. Also, when materials are still under copyright restrictions, they cannot be included in the Flickr Commons. It is a requirement for inclusion in the Commons that the images have no known copyright restrictions (Flickr from Yahoo, 2010d). Thus, copyright status determines the kind of actions an institution can take with their images, and it needs to be considered carefully.

The last point which was often made in the comments was that a Flickr collection requires a considerable amount of time in terms of planning, preparation, and management:

- “Establish standards for sizing, tags, and licensing/rights management. Have a photo release agreement ready for users ...”
- “[M]ake a list of everything you have before you decide to upload it to the web.”
- “I would warn other institutions, though, that sometimes it can be very time intensive to do all that tagging.”
- “I would recommend using a photo program that enables you to upload directly from your photo management system on the computers you are using.”

- “I would recommend that libraries be aware that their Flickr site is yet another source of comments and messages to monitor, in addition to whatever comes in from the library's website and/or blogs.”
- “I would recommend that they have a plan on how they want to use it before posting photos. And to monitor it regularly.”
- “Don't treat your Flickr site as a digital dumping ground. Clearly label pictures, sets, and collections. Be consistent. Include a link back to your organization's website.”
- “I recommend that an institution have a really detailed map of how they intend to group their various photos from the most basic levels to the larger groups available in the Flickr software.”
- “We started without clear goals or a plan. Making it up as we go along is ok, but it really should be a more organized effort.”

These comments show that in retrospect, many institutions feel that having clear goals and intentions before creating a Flickr collection makes maintaining the collection easier. In-depth preparation and concentrated management of the collection is, however, time consuming. Thus, if institutions create Flickr collections in order to increase the exposure of archival materials, they should be aware of costs involved in staff time dedicated to the project.

Factor Analysis Results

Using an alpha level of 0.05, the factors which influence indicators of success with statistical significance are as follows: age of the collection, whether or not social media features are enabled, time spent on metadata, and staff size. Of all of the factors, age is related to the greatest number of indicators of success. As age increases, the total number of views that a Flickr collection has received increases as well as the percentage of items which have received at least one comment. In addition, the time spent on reference questions in total increases as age increases.

The second factor which affects a few of the success indicators is the time spent on metadata. If the time spent on metadata is treated as a binary variable (either time is spent on metadata or it is not, meaning that metadata exists or does not), then metadata is related to an increase in both average monthly views and total views. Time spent on metadata also relates to an increase in total time spent on reference questions.

The third factor which has an effect on the success indicators is whether or not social media tools are enabled for the Flickr collection. Flickr allows account holders to disable and enable different tools, and the results of this study show that if the social functions are enabled, then the average monthly views and the total views increase.

Lastly, the number of staff members who dedicate time to the Flickr collection relates to a couple of the success indicators. As staff size increases, both the average monthly views and the total views for a Flickr collection increase.

Discussion

Determining the population for this study was challenging because a comprehensive list of digital archival collections on Flickr does not exist. There is also no tool within Flickr which allows account holders to identify themselves as such a collection on the Web site. The Flickr Commons is one attempt to create such a collocation, but it does not contain all of the digital archival collections on Flickr. The researcher had limited means for determining the population, so it is likely that the actual number of digital archival collections on Flickr is larger than the 207 which received the e-mail solicitation.

1a. What are the experiences of institutions which host digital archival collections on Flickr outside of the Commons in terms of usage statistics, goal achievement, and increased user awareness?

When the respondents were asked to choose the most important reasons for creating their Flickr collections, the three reasons rated the highest were to “expose collections to a broader audience / facilitate discovery of our materials,” “help advertise / provide a link to our institution,” and to “utilize Web 2.0 features to engage user involvement / discussion.” In addition, institutions chose the options “the photo is regionally important” and “the photo is part of a popular collection” most often as reasons why they chose these particular materials for hosting online. These responses

indicate an emphasis within these institutions on increasing user awareness of and participation with their materials, and Flickr is a part of achieving that goal.

Most of the institutions in this survey felt that they had accomplished their goals for their Flickr collections (80%). They also recommended Flickr and plan on continuing to add items to their collections there. Some of the common positive comments about Flickr included the fact that Flickr is easy to use, has low initial costs, and has an upgrade to a Pro account available at a low cost. A few institutions noted the price of the Pro account as a negative for Flickr, however, so the cost may be a burden on some institutions. A summary of the recommendations will be discussed in a later section.

While many of these institutions feel their goals have been met and state that they will continue to use Flickr, most of these collections have also received low usage statistics. An average of approximately 5.6% of these institutions' photos received at least one comment, 2.2% received at least one note, and 3.3% received at least one tag. Many institutions are not actually collecting these statistics or do not know them. If the average monthly views which members of the Commons receive (44,000) is an indication of the number of views that these collections could be receiving through Flickr, then their figures are quite low (Vaughan, 2010, p. 197).

Also, a surprising amount of institutions have the ways in which users can interact with the materials disabled. One of the highest rated reasons for using Flickr was to enable users to use Web 2.0 tools to interact with the materials; disabling the ability to interact seems counterproductive. Tags are disabled most often, which is not surprising since libraries and archives often prefer to use controlled vocabularies. One of the challenges with metadata in particular is creating a balance between standards and user

accessibility. One respondent advised: “when tagging, consider audiences other than the scholarly.” Another respondent noted that “[It] is difficult, though to bring a [F]lickr collection in line with ‘best practices’ for preservation and metadata,” which reveals the desire to have standards within the collection. Yet another concern, which was mentioned by multiple respondents and which applies to all of the possible user interactions, is the amount of time it consumes to monitor and moderate the user input. As it is now, however, users are not overwhelming many of the institutions with input. The decision comes down to the purpose of the Flickr collection and the amount of user interaction desired. As will be discussed later, institutions should be aware that when more functions for interaction are enabled, the amount of views a collection receives increases.

One of the most problematic circumstances this survey revealed is that many institutions are not collecting statistics which would allow them to evaluate the achievement of their goals. 38-59.5% of the respondents either did not respond or did not know the answers to the questions about average views per month, total views, and percent of items which have received at least one tag, one note, or one comment. Yet most of the institutions felt that they had met their goals for these collections. Since the institutions created these Flickr collections mainly to serve the users, it follows that they would record indicators such as the ones discussed in order to evaluate the success of the project. This study indicates that many do not.

Some respondents were notably aware of the low numbers as a concern or note that they should be tracking it:

- “As a smaller institution, the statistics have not been overwhelming, but I do like to see that items are being viewed and I am pleased whenever anyone adds one of our photos to their favorites or to a gallery. We have not had a lot of activity yet, but for a one-person endeavor, I think it's doing okay.”
- “I would do more promotion and publicity and invite people to enter into discussion to generate more user interaction.”
- “We would like to focus on how to increase our user participation. We had hoped to have much more user generated tagging and commenting on our photos, but have seen almost none.”
- “Your survey encouraged me to investigate how often our material on [Flickr] is being used.”
- “I plan to continue uploading content to Flickr. I have goals for the future, including uploading further photos, asking community members to comment and add information where I am lacking in it, and marketing (which has not been done yet - in that respect, I'm really quite happy with the number of hits we have had on photos so far).”

The recommendations section at the end of the paper will discuss some suggestions provided by the factor analysis and the comments from respondents. Those recommendations will address the concerns about usage.

1b. How do the experiences of the institutions in 1a differ (or not) from the experiences of institutions with collections in the Flickr Commons?

The institutions in the Flickr Commons (27 at the time of Vaughan's survey) chose the same top three reasons for creating their Flickr collections as the institutions outside of the Commons (Vaughan, 2010, p. 190). However, using Flickr in order to "utilize Web 2.0 features to engage user involvement / discussion" was the second most important reason for institutions in the Commons. This difference perhaps indicates their willingness to allow user generated tagging. Overall, though, institutions in and out of the Commons have similar goals which reflect the fact that while these institutions are diverse, their focus is on user awareness in whatever context.

When the participants were asked why they included these particular photos on Flickr, the results in this survey were opposite responses to Vaughan's study. The institutions in this survey chose materials for online access because they were in a popular collection and/or they were regionally important. Institutions in the Commons chose materials for the Flickr collection because they wanted to know what users could tell them about the photos and/or the photos were already online (Vaughan, 2010, p. 191). The difference revealed here may be in part because 64% of the institutions in the Commons also said that all of their materials were available online elsewhere (Vaughan, 2010, p. 197). Only 21% of the institutions outside of the Commons stated that all of their materials were available elsewhere. Institutions in the Commons might have the Flickr collection as a special project in order to take advantage of what the Commons in particular offers. In contrast, many institutions in this survey use the Flickr collection as the primary host for the collection because it is a way in which to have a digital archival collection within their budget constraints.

For both this study and Vaughan's study, the respondents indicated that on average they spent the most time on selecting and uploading photos (Vaughan, 2010, p. 192). One of the areas of significant difference was in time spent on monitoring and moderating user input. 28% of the respondents in this study spent no time that task, while none of the institutions in the Commons stated that they spent no time on that task (Vaughan, 2010, p. 192). A task which revealed a similar difference was time spent on responding to reference questions. 15% of the respondents in this survey spent no time on that task, while all of the institutions in the Commons spent at least some time on that task (Vaughan, 2010, p. 192). The differences here are likely rooted in the fact that the Commons receives so much promotion within Flickr and therefore receives more visitors.

One of the most significant differences between collections in the Commons and collections outside of the Commons is the amount of items which have received at least one tag, at least one comment, or at least one note. An average of approximately 46% of all Commons photos received at least one comment, 19% received at least one note, and 66% received at least one tag (Vaughan, 2010, p. 195). In contrast, an average of approximately 5.6% of the photos of the institutions in this survey received at least one comment, 2.2% received at least one note, and 3.3% received at least one tag. The wide gap here is again most likely because the Commons receives so much more publicity and promotion within Flickr. Also, all of the functions which allow users to interact with the photos are enabled.

As expected, the institutions in the Commons received higher views on average than the institutions in this study. Institutions in the Commons received an average of 44,000 views per month while the institutions in this survey received an average of

approximately 4,000 views per month (Vaughan, 2010, p. 197). This difference is likely, again, because of the much higher promotion the Commons received compared to the institutions which are independent of the Commons. Also, because many of the institutions included in the Commons are well known, they are perhaps more sought out by users and receive more views because of their prominence.

Overall, institutions in the Commons reported that they had met their goals. All but one respondent said they had met the majority of their goals, with the only other response being that they had met “some” of their goals (Vaughan, 2010, p. 198). In a similar vein, only one of the institutions in the Commons reported that the “popularity and impact” was less than expected (Vaughan, 2010, p. 198). Lastly, all of the respondents stated they would continue to use Flickr (Vaughan, 2010, p. 198), and all recommended it (Vaughan, 2010, p. 199). In contrast, 5% of the institutions in this survey reported that did not meet their goals for popularity and impact, 65% said they met some of their goals, and 31% said they met the majority of their goals. Some respondents in this survey also stated they would not recommend Flickr or that it would depend on the circumstances. The institutions in and out of the Commons had the same goals for increasing user awareness, but the institutions in the Commons received much more user activity. Since they had evidence of increased user awareness, they more clearly achieved their goals.

2. What are the factors which contribute to the “success” of a digital archival collection hosted on Flickr outside of the Commons?

The factors which affected the success indicators were age, time spent on metadata, the amount of functions enabled, and the number of staff involved in the Flickr collection. The first factor (age) unsurprisingly affected certain success indicators. Views and percentages increased as age increased since as time passes, it is more likely that a collection would have had some activity occurring. As the collection becomes well known and accessed by users more often, the time spent on reference questions would increase as well. The effect of this factor indicates that a Flickr collection needs time to be established and to begin receiving more user activity.

The second factor which affected success indicators is the time spent on metadata. Complete and descriptive metadata will bring more of the items into the results of a search, inform users about the items in the collection, and aid them in finding more items like it in the collection. The fact that metadata is related to an increase in views makes sense especially in terms of user searches. The items in a collection will not be retrieved as results in a search if there are no keywords attached to the item for which a user would search. When more users are finding and viewing the items, then more reference questions come to the institution regarding those items, which is the other success indicator with which metadata is associated. If the staff at an institution does not have the time to create metadata for the items, it might be a good idea to allow users to create metadata. The metadata would not be standardized, but it would bring up usage of the materials.

Another factor which related to a couple of the success indicators was the amount of social media tools which are enabled. When users are able to interact with the items and make use of social media tools which are becoming ubiquitous online, the amount of

views that a collection receives increases. Users have more interest in exploring the collection, sharing the content of the collection, and returning to the collection if it is more than a collection of static pages.

Lastly, the number of staff involved in the Flickr project is related to both the average monthly views and the total views for a Flickr collection. When more staff members work on the collection, the collection is likely better maintained simply because more time can be dedicated to it. With more staff members involved, the collection can be better monitored and more time can be spent uploading photos with complete metadata. If one staff member is given the daunting task of establishing a digital collection on Flickr alone, the time and thought necessary to make it a success will be missing.

Recommendations

Both the comments from respondents and the factor analysis suggest certain preparations to make before institutions create digital archival collections on Flickr.

Institutions which already have collections on Flickr can still review these recommendations and make adjustments as necessary and as applicable. Those recommendations are as follows:

1. Plan ahead. Select and organize photos ahead of time; do not treat the Flickr collection as a “digital dumping ground.”
2. Establish goals for the collection and collect statistics which will help in evaluation.
3. Dedicate an appropriate amount of staff time to the project in order to achieve your goals.
4. Create standards for the digitization process and the creation of metadata. Make sure the metadata is complete and thorough.
5. Know the copyright status of the photos you wish to upload and prepare policies.
6. Enable as many social media functions as possible in order to encourage user activity.
7. Purchase a Pro account, if possible, in order to track statistics and lose collection limits.
8. Give the project enough time to produce results.

Conclusion

This study has revealed that institutions with digital archival collections hosted on Flickr outside of the Commons have extra challenges if they are to achieve their goals. In the Commons, collections receive much more publicity and promotion, so institutions outside of the Commons need to be aware of the steps they can take to prepare their collections for success. The recommendations which are outlined in this paper were created by looking at statistical relationships through factor analysis as well as by finding common suggestions within the comments from respondents. Institutions which are considering Flickr as a host for a digital collection can use these recommendations as a starting point for the planning process and will be able to prepare their collection for success. Institutions which already have collections hosted on Flickr can use these recommendations to evaluate their collection and make adjustments which are evidence-based in order to increase user activity. In either case, the gathered and organized community knowledge in this paper will hopefully help institutions with new and current collections to achieve their project goals.

References

- Angus, E., Thelwall, M., & Stuart, D. (2008). General patterns of tag usage among university groups in Flickr. *Online Information Review*, 32(1): 89-101.
- Bond, T. J. (2006). Sustaining a digital collection after the grants: the Early Washington Maps Project. *OCLC Systems & Services: International digital library perspectives*, 22(1): 56-66.
- Flickr from Yahoo. (2010a). *Important update*. Retrieved from <http://www.flickr.com/commons/register/>
- Flickr from Yahoo. (2010b). *Participating institutions*. Retrieved from <http://www.flickr.com/commons/institutions/>
- Flickr from Yahoo. (2010c). *Upgrade*. Retrieved from <http://www.flickr.com/upgrade/>
- Flickr from Yahoo. (2010d). *Usage*. Retrieved from <http://www.flickr.com/commons/usage/>
- ILLINET. (2010). Appendix A-2011 for ILLINET/OCLC services users. Retrieved from http://www.cyberdriveillinois.com/departments/library/who_we_are/OCLC/pdfs/oclcpricing.pdf
- Kalfatovic, M. R., Kapsalis, E., Spiess, K. P., Van Camp, A., & Edson, M. (2008). Smithsonian team Flickr: A library, archives, and museums collaboration in web 2.0 space. *Archival Science*, 8(4): 267-277.
- Rafferty, P. & Hilderley, R. (2007). Flickr and democratic indexing: Dialogic approaches to indexing. *Aslib Proceedings: New Information Perspectives*, 59(4-5): 397-410.

- Samouelian, M. (2009). Embracing web 2.0: Archives and the newest generation of web applications. *The American Archivist*, 72(1): 42-71.
- Saunders, J. (2008). Flickr as a digital image collection host: A case study of the Alcuin Society. *Collection Management*, 33(4): 302-209.
- Stvilia, B. & Jorgensen, C. (2009). User-generated collection-level metadata in an online photo-sharing system. *Library & Information Science Research*, 31(1): 54-65.
- Vaughan, J. (2010). Insights into the Commons on Flickr. *portal: Libraries and the Academy*, 10(2): 185-214.

Appendix
Data Collection Instrument

Part I. Setting

1. At what type of institution are you employed?

- a. University library
- b. College library
- c. Public library
- d. Archive
- e. Museum
- f. Special library
- Other (please specify):

2. How large is the population which your institution serves (your institution's defined community)?

Part II. Background

1. What date did your first photo or collection 'go live' and become publicly viewable in Flickr? (month/year)

2. What were the primary reasons your institution was interested in joining Flickr? (Mark all that apply)

	Not applicable/not a consideration	Slightly important reason	Important reason	Very important reason
Expose collections to a broader audience / Facilitate discovery of our materials				
Utilize Web 2.0 features to engage user involvement / discussion				
We didn't already have these photos online and Flickr was a good system to initially publish them				
It would help advertise / provide a link to our institution				
Desire to join the company of other prestigious institutions that were already a part of Flickr				

Other (please specify and indicate how important the reasons was):

3. Specifically related to the selection of materials that you have incorporate into your Flickr presence, which factors below were important reasons you chose to incorporate that particular photo / collection into Flickr? (Mark all that apply)

- a. It was already available online
 - b. The photo is part of a popular collection
 - c. The photo is regionally important
 - d. We wanted to know more about the photo and wondered if viewers would be able to help provide information
- Other (please specify):

Part III. Institutional Staff Involvement

1. In total, how many individuals of any classification (staff members / volunteers / student assistants) at your institution are involved in the Flickr project on an ongoing

basis (e.g. selection, processing, uploading, technical development, answering questions from users, monitoring user comments/notes/tags, etc.)

2. Approximately how many hours per week, on average, are spent by your staff / volunteers / assistants on the following activities?

	None or not applicable	One hour or less	Between 1.1 and 5 hours	Between 5.1 and 10 hours	More than 10 hours
Selection of photos / collections with the direct intent that they will be incorporated into Flickr					
Adding metadata to photos with the direct intent that they will be incorporated into Flickr					
Uploading photos into Flickr					
Responding to reference type questions / user inquiries from users viewing your Flickr photos					
Conducting original research to substantiate / validate information provided by users viewing your Flickr photos					
Modifying records hosted in an existing local systems to reflect substantiated information that was provided by users					
Monitoring and / or moderating user generated content (notes, tags, comments)					
IT technical work (development of applications utilizing the Flickr API; development of different batch uploading/ingestion processes, etc.)					
General administrative / project management / future directions meetings related to Flickr					

Please list any other significant ongoing activities associated with your Flickr involvement, and approximately how many hours per week this activity:

Part IV. Marketing

1. If your institution made a concerted effort to market the Flickr photos collection, what methods were used? (Mark all that apply)

- a. Flyers / brochures / bookmarks
- b. E-mail notifications
- c. Signs / posters
- d. Access points on the institution's website
- e. Advertisements on the institution's website
- f. In orientations / classes

Other (please specify):

2. When did your institution begin implementing marketing for the Flickr photos / collection?

- a. More than 6 months before 'going live'
- b. 3-6 months before 'going live'
- c. 1-3 months before 'going live'
- d. At the point of 'going live'
- e. After 'going live'
- f. Not applicable

3. If your institution marketed the Flickr photos / collection, in what ways is the institution still marketing the photos / collection?

- a. Flyers / brochures / bookmarks
- b. E-mail notifications
- c. Signs / posters
- d. Access points on the institution's website
- e. Advertisements on the institution's website
- f. In orientations / classes

Other (please specify):

Part V. Social Interactions

1. Which types of social interaction / community building features have viewers used with your collections (regardless of the degree of use)?

	Yes, this feature has been used by some users	No, this feature is enabled but hasn't been used by any users at all	Not enabled	Don't know
User generated tags				
User generated comments				
User generated notes				
"Blog This" functionality				
Bookmarking of photos as "favorites"				
Incorporation of photos into viewing / discussion / subject groups set up by other Flickr members				

2. Approximately what percentage of your Flickr photos have received at least one tag by a viewer?

3. Approximately what percentage of your Flickr photos have received at least one comment by a viewer?

4. Approximately what percentage of your Flickr photos have received at least one note by a viewer?

Part VI. Statistics

1. Approximately how many views have you had on your overall collection of photos, in total, since launch?

2. Approximately how many views overall are you averaging per month at this time?

3. Approximately what percentage of photos currently in Flickr are also available through some other publicly accessible online venue (e.g. on a webpage that's part of your institution website, in a digital asset management system you or a collaborative host, etc.)

Part VII. Assessment

1. Have you noticed an increase in visitation to any local site resources (your institution's webpage, digital asset management system, etc.) that you feel is at least in part due to your involvement with Flickr?

a. Yes

b. No

2. In an earlier question, you selected some primary reasons your institution was interested in joining Flickr (such as exposing collections to a broader audience, utilizing web 2.0 features to engage user involvement / discussion, etc.). In general, do you feel that you've met your initial goals?

a. Yes

b. No

3. Somewhat related to the above question, how would YOU rate the overall popularity and impact of your Flickr photos / collections? "Overall popularity and impact" could refer to such factors as overall number of views for your photos, amount of user interaction as shown by user contributed tagging and comments, increased visitation to your institution's website, an increase in reference questions related to photos in the collection(s), etc. Please answer based not just on the activity generated at the initial launch, but rather, measured and averaged over time, from initial launch to the present.

a. The overall popularity and impact has exceeded what we would have expected.

b. The overall popularity and impact was about what we expected.

c. The overall popularity and impact was less than we expected.

4. Do you plan to continue your involvement with Flickr? If so, do you anticipate expanding the number of photos / collections you already have on Flickr? If you had to provide an estimate figure, approximately how many additional photos would you anticipate adding to the Flickr over the next twelve months?

5. Would you recommend becoming a member of Flickr to other institutions? Is there anything you would recommend they be especially aware of or should especially consider?

6. Additional comments?